



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE ECONOMÍA ♦ DIVISIÓN DE ESTUDIOS DE POSGRADO

PROGRAMA ÚNICO DE ESPECIALIZACIONES EN ECONOMÍA

*Determinantes del desempeño sobresaliente de estudiantes mexicanos de  
15 años en Competencia Lectora  
(PISA-2009, Análisis Logístico)*

## ENSAYO

QUE PARA OBTENER EL GRADO DE:  
Especialista en Econometría Aplicada

PRESENTA:  
Estephanie Santín Filloy

TUTOR:  
Dr. Miguel Cervantes Jiménez

CIUDAD.DE.MÉXICO , JUNIO DE 2017



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## **Resumen**

Con el propósito de analizar formalmente los determinantes para obtener un puntaje *sobresaliente* en la prueba PISA, se estima un modelo Logit que explica un puntaje superior a la media de la OCDE mediante las variables: género, edad en que el estudiante inicia los estudios de primaria, nivel de estudios de la madre, número de libros en el hogar y minutos que se destinan a lectura de ocio. Se utiliza información obtenida de la prueba PISA 2009, para competencia lectora.

Se plantean siete escenarios dónde –utilizando el coeficiente de odds y el cálculo de los efectos marginales– se encuentra qué, en México existe un impacto positivo en la probabilidad de calificar por encima del promedio en la prueba si: se cuenta en casa con más de 100 ejemplares de libros, el estudiante destina más de 30 minutos para lectura de ocio, el nivel de estudios de la madre es superior al bachillerato, el estudiante comenzó los estudios de primaria a una edad menor de los 7 años. Además, fue posible observar que existe una brecha de género en la habilidad lectora, con una ventaja para las mujeres.

Los dos últimos escenarios incluyen las variables que provocan un desempeño sobresaliente en competencia lectora y otro que provoca lo contrario, con los cuales se reafirman los hallazgos. Por último, se incluyen recomendaciones de política que buscan incidir positivamente en las variables analizadas para posicionar a más estudiantes como sobresalientes en el área competencia lectora.

### **Palabras clave**

Competencia Lectora, Educación, Logit, Prueba PISA 2009

### **Clasificación JEL**

I21, C25

## **Abstract**

The aim of this research is to analyze the determinants associated on obtaining an outstanding performance, i.e. a score above the OECD average on the PISA 2009 reading assessment, by Logistic regression analysis, using the following variables: gender, age when first entering school, mother's level of education, the number of books present at home and leisure time reading.

Seven scenarios are analyzed with odd ratio and marginal effects methodology. The main results are: there is a positive impact in the probability of scoring above the average of ODCE Pisa 2009 if: more than 100 books are present at home, the student reads more than 30 minutes for leisure, the mothers level of education is high school or higher, the student begun elementary school at 7 years old or less. An additional finding is that gender gap exists in reading assessment providing an advantage to girls.

The last two scenarios include in one hand, the group of variables that achieve the highest performance and in the other hand, the group of variables that obtain the lowest performance, on reading assessment. Finally, policy recommendations are included which would help to improve the performance of students on reading assessment.

## **Keywords**

Reading Assessment, Education, Logit, 2009 PISA test

## **JEL Classification**

I21, C25

## Índice

Introducción

1. Aspectos Teóricos.....	1
2. Estado de arte.....	3
3. Modelación econométrica.....	5
3.1 Modelo logístico.....	8
3.2 Forma del Modelo.....	9
4. Análisis y discusión de los resultados.....	12
4.1 Cálculo de probabilidad y generación de escenarios.....	17

Conclusiones

Recomendaciones

## Introducción

Un tema de interés en los estudios que buscan comprender la posición de los países en el desarrollo económico mundial es conocer, qué tan efectivo resulta el sistema educativo de los países; notables metodologías a lo largo del tiempo han tomado en cuenta el nivel de preparación de los ciudadanos para explicar fenómenos complejos como el desarrollo económico o la generación de ingresos en una sociedad, ciertamente se piensa que una sociedad bien preparada siempre será una sociedad próspera; por lo anterior se analiza el desempeño del sistema educativo mexicano, mediante la exploración de las características que determinan a los alumnos considerados como sobresalientes.

La experiencia mundial muestra la existencia de una estrecha relación entre el nivel de desarrollo de los países, en su sentido amplio, con la fortaleza de sus sistemas educativos, de investigación científica y tecnológica. El tema de la educación se identifica como determinante para el incremento en la movilidad social de los individuos, e indispensable para lograr avances científicos o tecnológicos, que demandan sociedades más preparadas ante el cambio, con los consecuentes efectos al enriquecimiento cultural y espiritual que provee a la humanidad: las sociedades que más han avanzado en lo económico y en lo social, son aquellas que han logrado cimentar su progreso en el conocimiento.

De acuerdo con diversos estudios de la Organización para la Cooperación y el Desarrollo Económicos (OCDE 2009a), un año adicional de escolaridad incrementa el PIB per cápita de un país entre 4 y 7 por ciento.

En este contexto, los resultados de la prueba PISA<sup>1</sup> han sido inquietantes para muchos países, la propia OCDE (2006) reconoce que el gran desafío planteado se encuentra en la laboriosa pero prometedora tarea de analizar a fondo los resultados de las evaluaciones, y no en la reacción momentánea que provoque para diseñar y poner en marcha políticas públicas.

---

<sup>1</sup> El Programa Internacional para la Evaluación de los Estudiantes es un estudio riguroso con aceptación mundial aplicado por la OCDE.

En este sentido, el objetivo del presente ensayo es identificar los factores que determinan el desempeño *sobresaliente*, es decir, por encima del promedio de la OCDE, en estudiantes mexicanos de 15 años para la prueba de competencia lectora. Para ello, se consideran los resultados que arrojó la prueba PISA 2009, analizados mediante un modelo de probabilidad logística (Logit).

Se espera demostrar que, si el nivel de escolaridad de la madre es superior al de bachillerato, si se destinan más de 30 minutos a lectura de ocio y si los hogares cuentan con más de 100 ejemplares de libros en casa, entonces, los estudiantes resultarán con una calificación sobresaliente en la prueba de competencia lectora. Asimismo, que existe una brecha, que bajo ciertas características podría favorecer al género femenino.

En la primera sección se señalan los aspectos teóricos, donde se aborda qué es la prueba PISA y cómo está conformada. En la segunda sección, se presenta el estado de arte, donde se exponen diversos hallazgos en el ámbito internacional a partir de la revisión de los resultados obtenidos en la prueba de competencia lectora. En el tercer apartado se explica la modelación econométrica, la cual se basa en un modelo Logit donde se analiza el coeficiente de odds y los efectos marginales. En la cuarta sección se muestran las conclusiones y, por último, se presentan las recomendaciones.

## 1. Aspectos teóricos

En este apartado se analiza que es la prueba PISA, su periodicidad y las áreas que la integran enfatizando el área de competencia lectora, que es el área en la que se centra el análisis.

El Programa Internacional para la Evaluación de los Estudiantes (PISA por sus siglas en inglés) es un estudio coordinado por la OCDE que realiza para conocer el estado general de los conocimientos básicos en los países asociados, de manera que permite contrastar el grado de desarrollo de los sistemas educativos.

El estudio se focaliza en estudiantes de 15 años de edad y se concentra en la evaluación de tres áreas: competencia lectora, competencia matemática y competencia científica. Mediante sus resultados se puede observar en qué medida los estudiantes son capaces de utilizar los conocimientos y habilidades adquiridas en la escuela para enfrentar los retos de la vida real, además de proporcionar los elementos necesarios para una evaluación de los contenidos considerados en los planes de estudio.

La participación en PISA ha sido extensa, ya que actualmente, participan todos los países miembros de la OCDE, así como varios países asociados, por lo que se puede considerar que sus resultados son de conocimiento y validez internacional.

Ahora bien, una de las áreas de interés en el presente estudio es lo que se ha denominado como competencia lectora, la cual se define como: la habilidad individual de entender, usar, reflexionar y vincularse con textos escritos<sup>2</sup>; dicha

---

<sup>2</sup> A pesar de que no existe una convención establecida para su traducción al español, este concepto proviene de la expresión inglesa "*literacy*", que en la acepción al español deriva de una amplia gama de definiciones, tales como: alfabetización, competencia, habilidad y cultura, incluso se ha llegado a emplear la palabra *literacidad* para explicarle. Independientemente de su traducción, para PISA es un concepto fundamental que se refiere, por un lado, a la capacidad para aplicar conocimientos y destrezas en situaciones diversas y, por otro, a la consecución de procesos cognitivos complejos, tales como analizar, razonar o comunicarse de manera efectiva; así como plantear, resolver e interpretar diferentes problemas. Como se aprecia, el concepto de alfabetización (o *literacy*) usado en PISA, va más allá de la idea tradicional de la capacidad de leer y escribir: es considerada como una habilidad que se aprende a lo largo de la vida y no como un rasgo que un individuo tiene o no tiene.



área cuenta con un apartado dedicado dentro del conjunto de apartados de la PISA, y se evalúa en relación a: un formato de texto, un proceso de lectura y diversas situaciones en que se expresa dicha lectura.

Para realizar el análisis del formato de texto, se toma en cuenta el rango de material escrito al que los estudiantes se enfrentarán más adelante de su vida; tanto textos continuos: narrativa, exposición y argumentación; como no continuos, por ejemplo: listas, formatos, gráficas o diagramas. Por su parte, en el proceso de lectura (también conocido como aspectos), se evalúa a los alumnos en su capacidad de mostrar eficiencia en obtener y recuperar conocimiento, así como su capacidad para formar un amplio conocimiento sobre un texto, interpretarlo y reflexionar su contenido, forma y características (OCDE 2009b). Por último, a las situaciones –que se definen por el uso para el cual fueron creados los textos– se les divide en cuatro categorías: 1) personal, como las novelas o una biografía; 2) pública, como documentos oficiales; 3) ocupacional, como un manual o un reporte, y; 4) educativa, como un libro de texto.

La prueba PISA se aplica cada tres años y en cada ciclo se enfatiza una de las tres áreas que evalúa: lectura, matemática o ciencias. En 2009 el área principal de evaluación fue Lectura, por lo cual se decidió tomar esta base de datos y no las subsecuentes de 2012 y 2015, que se centraron en preguntas relativas a la competencia de matemáticas y ciencias, respectivamente.

Los instrumentos de medición son cuadernillos y en ese ciclo se utilizaron 13 versiones, a cada estudiante se le asignó una sola versión. Los cuestionarios de contexto son otro tipo de instrumento que aporta información sobre el entorno de los estudiantes y las escuelas, y contribuye en el análisis y comprensión de los resultados. En PISA 2009 hubo seis cuestionarios de contexto: Cuestionario del Estudiante, Cuestionario Escolar, Cuestionario de Carrera Educativa, Cuestionario de Tecnologías de la Información y Comunicación, Cuestionario sobre la Lectura para la Escuela y Cuestionario para Padres.

En México se aplicaron los siguientes cuestionarios: Cuestionario del Estudiante, Cuestionario Escolar, Cuestionario de Carrera Educativa y

Cuestionario sobre la Lectura para la Escuela.

La prueba PISA 2009 se aplicó durante el mes de marzo de dicho año, a una muestra de 38,250 estudiantes en 1,535 escuelas secundarias o de educación media superior. La muestra de estudiantes es representativa a nivel nacional y para cada una de las 32 entidades federativas.

## **2. Estado del arte**

En este apartado se realiza una revisión a los hallazgos internacionales para las variables de análisis, con el propósito de conocer cómo incidieron en el puntaje de competencia lectora.

Una revisión a la literatura internacional nos ofrece una amplia gama de factores que afectan el desempeño de la competencia lectora en la prueba. Por ejemplo, Shiel y Eivers (2009) encuentran que uno de los factores que determina el desempeño de la lectura en la prueba PISA es el género. En este mismo sentido, Williams, T., Williams, K., Kastberg y Jocelyn (2005) encontraron que el género (a favor de las mujeres) influye tanto a la lectura como al desempeño en otras áreas de la prueba. De acuerdo al estudio de Topping (2006), utilizando la prueba de PISA, las mujeres tienen un desempeño superior al de los hombres en la prueba de competencia lectora, para todos los países.

Por su parte, el estudio de Sprietsma (2007), quien utiliza la información de PISA 2003, señala que un factor importante es la edad en que los estudiantes entran a la primaria. En dicho estudio se concluye que este factor tiene un impacto significativo en el largo plazo, encontrando que los alumnos que inician la primaria a más temprana edad obtienen una desviación estándar 20 por ciento inferior que aquellos que tenían una edad mayor. Por su parte, Suggate (2009), encontró que existe una correlación positiva entre el desempeño en competencia lectora y la edad en la que ingresan a la primaria los alumnos evaluados. Es por ello que se recomienda tener precaución al evaluar esta variable.

Otro factor relevante, manifiesto en diversos estudios relacionados con el tema, es el nivel de educación de los padres. Rasmussen (2003) concluye que existe una relación positiva entre estudiantes cuyas madres tienen educación universitaria y mayores puntajes en la prueba de lectura de sus hijos. En este mismo sentido, Geske y Ozola (2008) destacan que aquellos estudiantes con mayor puntaje en las pruebas de lectura están asociados a padres que destinan mayor tiempo a la lectura. Kasapoglu (2014) señala que, si se analiza la escolaridad del padre y de la madre respecto al desempeño en habilidad lectora, se halla que el número de años de estudio del padre es no significativo y el de la madre si lo es.

De acuerdo con algunas fuentes consultadas, el número de libros en el hogar de los estudiantes también es una variable que determina mejores niveles de desempeño en la prueba de lectura, Smith y Barret (2011) analizaron las características del desempeño en la prueba de lectura considerando dos grupos de países: el grupo "a" incluyó a Uganda, Kenia, Tanzania, Malawi, Mozambique y Zambia, caracterizados por ser países con ingresos nacionales bajos; el grupo "b" incluía Botswana, Namibia, Lesoto y Suazilandia cuyos ingresos nacionales son medios bajos. Los hallazgos de dicho estudio señalan que aquellos hogares que tenían más de cien libros mostraban mejor desempeño escolar que aquellos que no. Topping (2006) concluyó que los hogares de estudiantes que tenían más de cien libros obtenían mayor puntaje que aquellos que tenían menos de 10 libros. Por su parte, Hvistendahl y Roe (2004) hallaron que para el caso de Noruega esta variable no es significativa. Por lo cual esta variable debe de analizarse con cuidado.

La actitud hacia la lectura también ha sido identificada como un factor que incide en el desempeño. Hvistendahl y Ro (2004) señalan que existe una relación positiva para el caso noruego; sin embargo, también es una variable que debe tratarse con cuidado ya que, para Twist, Gnaldi, Shagen y Morrison (2004) la variable resultó negativa sobre el desempeño escolar de los estudiantes en Inglaterra.

Por último, un buen desempeño en la lectura se asocia frecuentemente a la cantidad de tiempo que se lee en casa. Se ha encontrado una fuerte correlación entre el tiempo que se invierte en leer por ocio y la brecha entre alto desempeño y bajo desempeño de la competencia lectora (Topping 2006). Incluso Brozo, W. G., Shiel, G., & Topping, K. (2007). sugiere que es recomendable incrementar el tiempo de lectura en casa, ya que tendrá resultados positivos en el desempeño de los alumnos en el ámbito de la lectura en países como Estados Unidos y Reino Unido.

### **3. Modelación econométrica**

En este apartado se describen los datos a emplear, también se desarrolla y justifica la selección de una herramienta estadística adecuada para su análisis, que nos permita contrastar la hipótesis planteada.

Para el análisis, se utiliza la base de PISA 2009 para México, obtenida de la página web de la OCDE, particularmente la información que arroja el Cuestionario del Estudiante y el puntaje de capacidad lectora.

Tomando en cuenta lo anterior, se realizaron modificaciones a las variables, para estar en condiciones de realizar un análisis más puntual de la información, de la siguiente forma:

El promedio de competencia lectora de la OCDE en la prueba PISA 2009 es de 493 puntos, por ello se transforma dicha variable de la siguiente manera: 1 cuando el promedio de competencia lectora del estudiante es igual o mayor a 493 y 0 cuando el promedio de competencia lectora del estudiante es menor a 493.

Las variables predictoras se dividen en dos grupos:

- a) Variables de contexto, que incluye al género, edad en la que el estudiante entró a primaria, el grado de estudio de la madre y número de libros en casa.
- b) Variable de tiempo, que incluye una variable relativa al tiempo de lectura

por ocio de los estudiantes.

La variable de género, se incluyó sin transformaciones, donde 1 es género femenino y 2 masculino.

La variable de edad en la que el estudiante entró a primaria tiene valores que van de 4 a 10 años, es decir la edad en la que los alumnos comenzaron sus estudios de primaria, el análisis se reduce a dos casos, menores a 7 años (igual a cero) y de 7 años en adelante (igual a 1).

La variable que refiere al nivel máximo de estudios de la madre se divide en cinco categorías: 5 si no terminó la primaria, 4 si terminó la primaria, 3 si terminó la secundaria, 2 si tiene estudios técnicos y 1 si tiene estudio bachillerato o mayores; para simplificar el análisis, se divide la variable en dos categorías: 1 si tiene estudios de bachillerato o mayores y cero si tiene estudios técnicos o inferiores.

Respecto a la variable referente al número de libros que hay en casa del estudiante se divide en 6 categorías:

**Tabla 1. Número de libros en casa del estudiante**

Número de la Categoría	Número de libros que hay en casa del estudiante
1	0 a 10
2	11 a 25
3	26 a 100
4	101 a 200
5	201 a 500
6	500 o más

Fuente: elaboración propia con base en OCDE

Con el propósito de caracterizar a esta variable como categórica binaria, se dividirá la variable en dos secciones, de 0 a 100 libros (igual a 1) y 100 o más libros (igual a 0).

Por último, la variable referente al número de minutos que los estudiantes

dedican a leer por gusto se divide en cinco categorías:

**Tabla 2. Número de minutos dedicados a leer por gusto**

Número de la Categoría	Número de minutos dedicados a leer por gusto
1	No leen por gusto
2	Lee menos de 30 minutos
3	Lee más de 30 minutos, pero menos de 60 minutos
4	Lee de 1 a dos horas por día
5	Lee más de 2 horas al día

Fuente: elaboración propia con base en OCDE

Para facilitar nuestro análisis y ser consistente con el tratamiento de las variables anteriores, se dividirá en dos valores: 1 para aquellos que leen más de 30 minutos y 0 para aquellos que leen menos de 30 minutos incluso aquellos que no leen.

Para observar nítidamente la composición de los datos, se presenta la siguiente tabla:

**Tabla 3. Proporción de las variables**

Variable	Cualidad	Proporción
Genero	Hombre	53%
	Mujer	47%
Edad de ingreso a la primaria	<7 años	83%
	>7 años	17%
Nivel de estudios de la madre	<Media superior	64%
	>Media superior	36%
Cantidad de libros en casa	<100	90%
	>100	10%
Lectura de ocio	<30 min	69%
	>30 min	31%

Fuente: Elaboración propia con datos de la OCDE, 2009.

En la Tabla 3, se observa que el 53 por ciento de los estudiantes en la muestra son hombres y el 47 por ciento son mujeres. El 83 por ciento de los estudiantes ingresan a la primaria a una edad menor a siete años. Por su parte, el 64 por ciento de las madres de los estudiantes tienen un nivel de estudios menor a

bachillerato. Únicamente el 10 por ciento de los hogares de los estudiantes cuenta con más de 100 ejemplares de libros. Por último, el 69 por ciento de los estudiantes destinan menos de 30 minutos a la lectura por ocio.

### **3.1 Modelo logístico**

Durante la década de los setenta, se gestó un importante desarrollo en los modelos de elección discreta –o de respuesta cualitativa– caracterizado por su aplicación en las ciencias sociales, particularmente como una herramienta estadística que permitiera modelar los determinantes en el comportamiento de los agentes, según decisiones y contextos particulares. Si bien este tipo de modelación se empleó en años anteriores, principalmente para aplicaciones biométricas, fue hasta los trabajos de Heckman y Willis (1977) y McFadden (1976) cuando se comenzó a aplicar en las decisiones económicas y de política pública.

Según Amemiya (1985) se puede definir a un modelo logístico como un modelo de elección discreta, que cuenta con una función cuya variable dependiente toma la forma de una variable aleatoria dicotómica, donde puede tomar el valor de 1 si el evento ocurre o bien el valor de 0 si el evento no ocurre; teóricamente se asume que la probabilidad de ocurrencia del evento depende de un vector de variables independientes que multiplica a un vector de parámetros desconocidos.

El conjunto de regresores son las variables que permiten discriminar entre los grupos, y que determinan la pertenencia de un elemento a un grupo u otro, estos pueden estar medidos en escala nominal, ordinal, de intervalo o de razón.

El resultado del análisis es un vector de parámetros con valores numéricos, (coeficientes) para cada una de las variables explicativas que forman el modelo, al tenerse en cuenta todas en conjunto y dar valores a cada una de las variables independientes contenidas en el modelo definitivo, se obtiene el valor de la probabilidad de que un individuo posea la característica de interés estudiada en el modelo.

Además de obtener estimaciones de la probabilidad de un suceso, este modelo permite identificar los factores de riesgo que inciden sobre dichas probabilidades, así como la influencia o peso relativo que estos tienen sobre las mismas.

El resultado es un índice, que asociado con algún método de estratificación permite generar clasificaciones sobre los valores observados, a dichas agrupaciones se le asocia una calificación. Existen muchos criterios para llevar a cabo la asociación índice - calificación, desde aquellos puramente estadísticos (con base en índices de muestreo), hasta en los que el criterio podría considerarse como subjetivo, por ejemplo, los modelos que toman en cuenta características de los sujetos, como la preferencia de un artículo sobre otro.

Para el caso más sencillo, que comúnmente se ha denominado como modelo lineal de probabilidad, se asume que la relación entre las variables explicativas y la variable explicada tiene un comportamiento lineal, suposición que en muchos casos no se da, dando esta situación origen a los modelos de regresión no lineales, dentro de los cuales se encuentran ubicados los modelos Probit y Logit, siendo este último el que interesa en el presente análisis, por lo que se detallará su estructura y los fundamentos teóricos que lo sustentan.

La modelización Logit es similar a la regresión tradicional salvo que utiliza como función de estimación la función logística, por lo que el resultado del modelo es la estimación de la probabilidad de que un nuevo individuo pertenezca a un grupo o a otro; al tratarse de un análisis de regresión, también permite identificar las variables más importantes que explican las diferencias entre grupos.

### **3.2 Forma del Modelo**

Para el presente modelo se determina que la variable explicada (Y) es una variable dicotómica, que puede tomar solamente dos valores: el valor 1 representa un estudiante con puntaje en competencia lectora superior al promedio de la OCDE, mientras que el 0 representa un estudiante con puntaje



en competencia lectora inferior al promedio de la OCDE.

Se estimará por medio del siguiente modelo no lineal:

Siguiendo a Greene y Hensher (2009), para estimar la probabilidad de que un individuo posea una calificación por encima del promedio para los países miembros de la OCDE en competencia lectora se tiene:

$$Y_i = \frac{1}{1 + \lambda^{-\alpha - \beta_k X_{ki}}} + \varepsilon_i = \frac{\lambda^{\alpha + \beta_k X_{ki}}}{1 + \lambda^{\alpha + \beta_k X_{ki}}} + \varepsilon_i$$

Dónde:

$$M_i + M_i e^{\alpha + \beta_k X_{ki}} = e^{\alpha + \beta_k X_{ki}}$$

$$M_i = (1 - M_i) e^{\alpha + \beta_k X_{ki}}$$

$$\frac{M_i}{(1 - M_i)} = e^{\alpha + \beta_k X_{ki}}$$

El modelo logístico descrito será estimado por máxima verosimilitud, en donde se supone independencia entre las observaciones, la densidad conjunta de la  $Y_i$  será:

$$Prob(Y_1 Y_2 \dots Y_i \dots Y_n) = \prod_{i=1}^n M_i^{Y_i} (1 - M_i)^{1 - Y_i}$$

Donde  $M_i$  es la probabilidad de que ocurra el evento  $i$ , y ya se encuentra definida; si realizamos una transformación logarítmica de la función de densidad conjunta se tiene:

$$\mathcal{L} = \ln L = \sum_{i=1}^i Y_i \ln M_i + \sum_{i=1}^i (1 - Y_i) \ln(1 - M_i) = \sum Y_i \ln M_i + \sum (1 - Y_i) \ln(1 - M_i)$$

Dado que el objetivo de esta estimación es maximizar la función  $\mathcal{L}$ , se realizan condiciones de primer orden con respecto a los parámetros que se estiman y se resuelve el sistema:

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \sum_{i=1}^i (Y_i - M_i) = \sum Y_i - \left( \frac{e^{\alpha + \beta X_1}}{1 + e^{\alpha + \beta X_1}} \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = \sum_{i=1}^i (Y_i - M_i) X_i = \sum Y_i X_i - \left( \frac{e^{\alpha + \beta X_1}}{1 + e^{\alpha + \beta X_1}} \right) X_i = 0$$

Bajo supuestos generales, los estimadores obtenidos deberán ser consistentes, asintóticamente eficientes y con una distribución normal, por lo cual se aplican diferentes pruebas para garantizar estos supuestos.

A fin de observar que variables en el modelo no son importantes para determinar cambios en la variable dependiente se debe calcular el estadístico de Wald (1939), debemos tener en cuenta la escala de medición de la variable, ya que pueden ocurrir dos situaciones: que sea categórica o no categórica, para cada uno de los cuales se tendría lo siguiente:

$$Wald_i = \frac{\beta_1^2}{\sigma_{\beta_i}^2}$$

Si la variable es categórica, se tiene que sigue una distribución chi cuadrado, con grados de libertad igual al número de parámetros estimados.

$$Wald_i = \beta_i^T \beta_i^{-1} \beta_i$$

La importancia del estadístico de Wald radica en que, a través de dicho valor se puede determinar la significancia del parámetro en el modelo, así como el peso que tiene, lo que se logra al establecer una sencilla prueba de hipótesis en la que se plantea lo siguiente:

*Ho:  $\beta_i = 0$  La variable  $X_i$ , no es importante para establecer variaciones en el modelo*

*Ha:  $\beta_i \neq 0$  La variable  $X_i$ , si es importante para establecer variaciones en el modelo*

De manera general, el modelo supone homocedasticidad en los errores,  $\varepsilon_i \sim N(0, \sigma_i^2)$ , si se viola dicho supuesto, derivaría en un importante sesgo en los coeficientes estimados y por tanto pérdida de eficiencia.

Una forma de probar la existencia de heterocedasticidad es la propuesta por Harvey (1976), modelando la varianza en función del conjunto de variables, tal que:  $\sigma_i^2 = [e(z' i \pi \gamma)]^2$ , se observa que, si ninguna de las variables tiene un efecto en la varianza, entonces,  $\sigma_i^2 = e(0) = 1$

Sabiendo que los estimadores son consistentes y asintóticamente eficientes, se procede a calcular la probabilidad que tiene un individuo de tener un puntaje sobresaliente al promedio de la OCDE en competencia lectora, después se calculan los efectos marginales, es decir el efecto de la k-ésima variable explicativa, manteniendo el resto constante, lo cual podría expresarse como:

$$\frac{\partial \Pr(Y_i = 1 | x_i)}{\partial x_{ik}} = \frac{\partial F(x' \beta)}{\partial x_{ik}} = f(x_i' \beta) \beta_k$$

#### 4. Análisis y discusión de los resultados

Una vez planteado el modelo, se establecen las variables como categóricas y se realiza un modelo lineal generalizado, indicando que los factores que determinan alumnos cuya probabilidad de obtener un puntaje superior al promedio de la OCDE en competencia lectora sea más alto son: la edad en la que el alumno ingresó a la primaria, el nivel de escolaridad de la madre, el número de libros en casa, su género y el tiempo que se invierte en lectura por gusto.

**Tabla 4. Regresión logística**

<b>Variables independientes</b>	<b>Coef.</b>	<b>Err. Std</b>
Genero	-0.39*	0.027
Edad primaria	-0.59*	0.040
Grado de estudios madre	0.98*	0.027
Libros en casa	-0.66*	0.038
Lectura por ocio	0.43*	0.027
Constante	-0.51*	0.058
LR chi2	2933.53	
Prob>Chi2	0	
McFadden's R2	0.0764	
Clasificados correctamente	77.29%	
Observaciones	35311	

\*/ Coeficientes significativos al 1%

Fuente: elaboración propia con datos de la estimación.

Se observa que los coeficientes, los errores estándar, los z-estadísticos, así como los p-values asociados son significativos para los valores 1 en las variables, edad en la que el alumno ingresó a la primaria, el nivel de escolaridad de la madre, el número de libros en casa y el tiempo que se invierte en lectura por gusto; asimismo, dichos valores son significativos para la variable género cuando ésta toma el valor 2.

**Tabla 5. Regresión logística con errores estándar robustos**

<b>Variables independientes</b>	<b>Coef.</b>	<b>Err. Std</b>
Genero	-0.39*	0.027
Edad primaria	-0.59*	0.400
Grado de estudios madre	0.98*	0.026
Libros en casa	-0.66*	0.037
Lectura por ocio	0.43*	0.027
Constante	-0.51*	0.057
LR chi2	2714.74	
Prob>Chi2	0	
McFadden's R2	0.0764	
Clasificados correctamente	77.29%	
Observaciones	35311	

\*/ Coeficientes significativos al 1%

Fuente: Elaboración propia con datos de la estimación.

Con el propósito de conocer si el modelo tiene problemas de heteroscedasticidad (lo cual incidiría en una incorrecta estimación del nivel de los coeficientes), se calcula el mismo modelo, pero con errores robustos; al compararlos se observa que los errores estándar de ambos son muy similares, por lo cual se puede descartar que el modelo presente heteroscedasticidad, por tanto, se utiliza el modelo simple.

Para confirmar lo anterior, se toma en cuenta a la variable género como el vector de variables Z; para poder probar la existencia de homocedasticidad en el modelo, se obtuvo la siguiente estimación:

**Tabla 6. Prueba para heterocedasticidad**

<b>Variables independientes</b>	<b>Coef.</b>	<b>Err. Std</b>
Género	-0.2*	0.027
Edad primaria	-0.3*	0.400
Grado de estudios madre	0.52*	0.026
Libros en casa	-0.35*	0.037
Lectura por ocio	0.36*	0.027
ConstanteE	-0.37*	0.057
LR chi2	3352.67	
Prob>Chi2	0	
Insigma2 lectura por ocio	-0.25	0.04
Observaciones	35311	
Insigma2=0		
chi2	39.43	
Prob>Chi2	0	

\*/ Coeficientes significativos al 1%

Fuente: elaboración propia con datos de la estimación.

El resultado de la prueba LR indica que no existe una mejora significativa en el modelo resultado de una generalización a un modelo homoscedástico, por lo tanto, se acepta la hipótesis nula de homocedasticidad: se puede afirmar que el modelo proporciona estimadores insesgados y consistentes.

El estadístico pseudo  $r$  cuadrada calcula una razón de verosimilitudes, la cual indica el nivel de mejora que tiene el modelo completo respecto al modelo que considera sólo el intercepto.

Se encuentra que, el porcentaje de los valores correctamente especificados para ambas regresiones fue de 77.3 por ciento.

Considerando una muestra con 35,311 valores, se observa que las observaciones incorrectamente clasificadas con 1 cuando su correcta clasificación es 0 son 648; por su parte, 7,370 valores están mal clasificados como 0 cuando su valor correcto debe ser 1, las restantes están correctamente especificadas.

La medida de sensibilidad señala que, para  $y=1$ , el 10.6 por ciento del total de observaciones están correctamente especificadas, por su parte, la porción de observaciones correctamente especificadas como  $y=0$  es de 97.6 por ciento.

Por último, las razones  $\frac{648}{27065}$  y  $\frac{876}{8246}$  son conocidas como las tasas de clasificación incorrecta de falsos positivos y falsos negativos.

Se realiza la prueba Wald o de Chi cuadrada:

**Tabla 7. Prueba de Wald**

<b>X<sup>2</sup></b>	<b>GL</b>	<b>P-value</b>
2535.7	4	0.0000

Fuente: Elaboración propia con datos de la estimación.

El valor del estadístico chi cuadrada de 2,535.7, con 4 grados de libertad, está asociado a un p-value de 0.0 lo cual indica que el efecto del puntaje superior a la media de los países que integran la OCDE, así como países miembros es estadísticamente significativo para todas las variables que integran el modelo.

**Tabla 8. Regresión logístic con odds ratio**

<b>VARIABLES INDEPENDIENTES</b>	<b>Odds Ratio</b>	<b>Err. Std</b>
Genero	0.6763*	0.6763
Edad primaria	0.5507*	0.5507
Grado de estudios madre	0.98*	2.6695
Libros en casa	-0.66*	0.5129
Lectura por ocio	1.5409*	0.042
Constante	0.5974*	0.034
LR chi2	2933.53	
Prob>Chi2	0	
McFadden's R2	0.0764	
Clasificados correctamente	77.29%	
Observaciones	35311	

Fuente: Elaboración propia con datos de la estimación

El cálculo del cociente de odds facilita la interpretación de los parámetros

estimados si se quiere conocer qué tanto más probable es, que un cierto evento (en este caso que la variación en la preferencia o ventaja de un individuo cuando incrementa en una unidad alguna de las variables explicativas) ocurra en un grupo, respecto a su ocurrencia en otro grupo. Por ello, que se decide obtener el resultado de la regresión con esta forma e interpretar los resultados; a continuación, se presentan los hallazgos:

Se encontró que, los estudiantes que comenzaron la primaria con 7 años o más, respecto de iniciar a una edad menor, reduce la probabilidad logarítmica de obtener un puntaje mayor al promedio en lectura en 0.59 .

Por su parte, el hecho de que la madre del estudiante cuente con estudios de bachillerato o superiores, respecto a la situación contraria, cambia la probabilidad logarítmica de obtener un puntaje mayor al promedio en lectura 0.98 .

En relación al número de libros en el hogar, se comprueba que si el alumno tiene de 0 a 100 libros, respecto a tener más de 100 libros, la probabilidad logarítmica de obtener un puntaje mayor al promedio en lectura se reduce en 0.66 .

Los resultados obtenidos para la variable de género muestran que, si el alumno es hombre, respecto a si es mujer, la probabilidad logarítmica de obtener un puntaje mayor al promedio en lectura se reduce 0.39 .

La variable, tiempo que los alumnos invierten en leer por ocio, demuestra que, si el alumno invierte más de dicho periodo en leer por gusto, respecto a aquellos que no leen o leen menos, cambia la probabilidad logarítmica de obtener un puntaje mayor al promedio en lectura en 0.43 .

#### **4.1 Cálculo de probabilidad y generación de escenarios**

Para calcular la probabilidad de ocurrencia del evento, es necesario sustituir los coeficientes estimados en un escenario particular, que será definido por las variables dicotómicas del modelo, ya que –dada su composición– sería estéril



el realizar un análisis con las medias.

En un primer escenario se analiza la incidencia del género en la probabilidad de obtener un puntaje promedio mayor en la prueba de lectura:

**Tabla 9. Escenario de Género**

Pr (Y=1 x):	0.3195				
Escenario: cuando					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	1	1	1	1	1

Pr (Y=1 x):	0.241				
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	1	1	1	1	1

Fuente: Elaboración propia con datos de la estimación.

Se observa que, en igualdad de circunstancias, una mujer posee mayor probabilidad de ser sobresaliente en la prueba de lectura de la OCDE, con un 31.9 por ciento, mientras que ser hombre, reduce la probabilidad de ser sobresaliente pues en las mismas condiciones se tiene un 24.1 por ciento.

Otro análisis de interés resulta el planteado por un escenario donde todos los individuos de la prueba realizada hayan ingresado a primaria antes de los siete años de edad:

**Tabla 10. Escenario edad de ingreso a primaria**

Pr (Y=1 x):		0.4602			
Escenario:					
Género	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	0	1	1	1	1
Pr (Y=1 x):		0.3657			
Escenario:					
Género	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	0	1	1	1	1

Fuente: Elaboración propia con datos de la estimación.

Los datos muestran que, el ingresar a primaria antes de los 7 años es un factor importante para el éxito del experimento: para hombres y mujeres incrementa la probabilidad de ser sobresaliente en la prueba en 36.5 y 46 por ciento, respectivamente.

El siguiente escenario consiste en simular el impacto del nivel educativo de la madre en los resultados de la prueba, para ello se calcula la probabilidad de éxito dado un escenario en donde el nivel educativo de la madre es inferior al bachillerato:

**Tabla 11. Escenario grado de estudios de la madre**

Pr (Y=1 x):		0.1496			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	1	0	1	1	
Pr (Y=1 x):		0.1063			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	1	0	1	1	

Fuente: Elaboración propia con datos de la estimación.

El resultado demuestra que la probabilidad de ser sobresaliente en la prueba para hombres y mujeres se ve considerablemente mermada por la educación de la madre, si se simula que se tiene una educación por debajo de la media superior, la probabilidad cae a 14.9 por ciento de éxito para mujeres y 10.6 para hombres.

Se cree que otro factor de alto impacto para determinar los resultados en la prueba es la cantidad de libros con que cuenta en el hogar cada uno de los individuos, para demostrarlo se genera un escenario en donde los alumnos cuentan con más de cien ejemplares disponibles en el hogar.

**Tabla 12. Escenario más de 100 libros en casa**

Pr (Y=1 x):		0.4779			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	1	1	0	1	
Pr (Y=1 x):		0.3824			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	1	1	0	1	

Fuente: Elaboración propia con datos de la estimación.

Los resultados de este escenario comprueban que existe un impacto positivo en la probabilidad de calificar por encima del promedio en la prueba realizada por la OCDE si se cuenta con más de cien ejemplares de libros en casa.

Dentro de este modelo también se comprueba que el tiempo dedicado a la lectura por los individuos que participan en esta muestra es relevante para determinar su puntaje, por lo tanto, se generará un escenario en donde se restringe a los individuos con una lectura de menos de treinta minutos al día.

**Tabla 13. Escenario con tiempo de lectura menor a 30 min al día**

Pr (Y=1 x):		0.2336			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	1	1	1	0	
Pr (Y=1 x):		0.1709			
Escenario:					
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	1	1	1	0	

Fuente: Elaboración propia con datos de la estimación

Como último gran resultado en la medición de la probabilidad de éxito del experimento realizado, se observa que los datos sugieren que leer menos de 30 min al día reduce la probabilidad de ser sobresaliente en la prueba de lectura, para el caso de las mujeres la probabilidad de éxito –manteniendo las demás variables constantes– se reduce a 23.6 por ciento, y a 17 en el caso de los hombres.

Adicionalmente, se generan dos escenarios relevantes: el que nos acerca a la ocurrencia del evento, es decir a que el individuo que presenta la prueba tenga la mayor probabilidad de estar por encima del promedio de la OCDE; y un escenario en donde sea la mayor probabilidad de fracaso, es decir, que se esté por debajo o en el promedio:

**Tabla 14. Mejor escenario propuesto**

Pr (Y=1 x):	0.6244				
	Escenario:				
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
1	0	1	0	1	

Fuente: Elaboración propia con datos de la estimación

La Tabla 14 muestra que, según las variables definidas, los factores determinantes para estar por encima del promedio son: entrar a la primaria antes de los 7 años, que la madre cuente al menos con educación media superior, que posea más de cien libros en casa y se destine un tiempo superior a los treinta minutos diarios a la lectura.

Siguiendo las recomendaciones contrarias a la especificación del escenario anterior se encuentra lo siguiente:

**Tabla 15. Peor escenario propuesto**

Pr (Y=1 x):	0.0717				
	Escenario:				
Genero	Edad Primaria	Grado de estudios madre	Libros en casa	Lectura por ocio	
2	1	0	1	0	

Fuente: Elaboración propia con datos de la estimación.

Los resultados son contrastantes, ya que van del 62.4 al 7.2 por ciento en la probabilidad de éxito.

Para terminar el análisis de datos se revisan los efectos marginales existentes para el escenario de la probabilidad más alta:

**Tabla 16. Efectos Marginales**

Variables independientes	$dy/dx$	Z	Err. Std	X
Genero	-0.091**	-14.88	0.0061	1
Edad primaria	-0.1464*	-14.52	0.01	0
Grado de estudios madre	0.2406*	37.78	0.006	1
Libros en casa	-0.1641*	-17.96	0.0091	0
Lectura por ocio	0.1054*	15.64	0.0067	1

\*/  $dy/dx$  es un cambio discreto de la variable dicotómica de 0 a 1

\*\*/  $dy/dx$  es un cambio discreto de la variable dicotómica de 1 a 2

Fuente: Elaboración propia con datos de la estimación.

Los efectos marginales son un aporte poderoso, ya que indican la contribución que tiene un cambio discreto en las variables con respecto a la probabilidad de ocurrencia del evento.

De los resultados obtenidos se puede saber que la probabilidad de ser sobresaliente en la prueba si se es hombre cae en 9.1 por ciento respecto de ser mujer; el ingresar a la primaria a los siete años o después, disminuye la probabilidad de estar por encima de la media de la OCDE en 14.6 por ciento; que la madre del individuo posea estudios al menos al nivel de media superior incrementa la probabilidad de éxito en 24.1 por ciento; el tener menos de cien libros en el hogar la disminuye en 16.4 y leer más de 30 minutos al día la incrementa en 10.5 por cien

## Conclusiones

Luego de analizar las variables que mejor explican la probabilidad de obtener un puntaje sobresaliente en competencia lectora –de acuerdo con el marco metodológico de la Prueba PISA para el caso mexicano–, ha sido posible determinar que la edad en la cual se inicia la primaria resulta ser muy explicativa en el desempeño escolar, ya que tienen ventaja los alumnos que comenzaron dicho estudio a una edad menor de 7 años. Es interesante que la evidencia empírica para países como Alemania demuestra que incide positivamente en la habilidad lectora comenzar a estudiar la primaria a una edad de 7 años o mayor porque en dicho país se invierte una mayor cantidad de tiempo y recursos en la educación inicial.

Ha sido posible constatar que, en la medida que el nivel de estudios de la madre es mayor, más probable será que su hijo obtenga un mejor desempeño en la prueba. En este caso, el resultado mexicano coincide con la evidencia empírica internacional, lo que puede justificar la implementación de políticas que garanticen el acceso de las mujeres a la educación profesional (tal y como se aplica en aquellos países donde dicha variable obtuvo los resultados más altos).

Este argumento se refuerza con el análisis de la variable género, ya que fue posible observar que existe una brecha en la habilidad lectora, con una ventaja para las mujeres (resultados que fueron consistentes en casi todos los países en los que se realiza la prueba); por tanto, reconocer que un sistema educativo que busque sociedades equitativas en términos de la enseñanza básica y media superior en mujeres, debería garantizar que –a la postre– los alumnos de ambos sexos cuenten con mejores elementos para la obtención de conocimiento en a través de la lectura, superior al promedio actual.

En relación con el punto anterior, se observa que existe una relación positiva entre número de libros en casa y el desempeño en la prueba de lectura (tanto para hombres como para mujeres), lo cual coincide con el análisis internacional



donde se menciona que en países con ingresos medios a bajos existe una relación positiva entre desempeño escolar en lectura y número de libros en casa, contrario a las economías con ingresos altos, donde esta relación resulta no significativa.

De ello se desprende que, en México es necesario fomentar la adquisición de libros en hogares que tienen hijos que cursan la escuela primaria, y de preferencia si son niños, ya que ello determinará que adquieran un sentido de gusto por la lectura, que conlleva inequívocamente a mejores alumnos. Al respecto, el número de libros en casa, está fuertemente relacionado con el nivel educativo de la madre por lo cual se confirma la impresión mencionada en párrafos anteriores, que el recurso de apostar por una educación superior en el género femenino elevaría los estándares de conocimiento en ambos géneros.

En este mismo sentido se confirma –dado el estudio de la variable asociada al tiempo desinado a la lectura de esparcimiento u ocio– que los estudiantes que destinan 30 minutos o más en leer por su cuenta, ajena a alguna tarea escolar y de manera periódica, muestran un mejor desempeño en la prueba de lectura, lo que –en síntesis del estudio implementado a partir de la prueba Pisa– contribuye fuertemente a contar con individuos que puedan utilizar los conocimientos y habilidades adquiridas en la escuela para enfrentar los retos de la vida real.

## **Recomendaciones**

Este estudio identifica cuales son los factores que contribuyen a que un estudiante de 15 años sea sobresaliente en la prueba de competencia lectora en México, en este sentido, se señalan recomendaciones para lograrlo.

Dado que se observa que, a mayor tiempo destinado a la lectura de ocio se incrementa la probabilidad de ser un estudiante sobresaliente en competencia lectora, es posible permitir y fomentar en los alumnos elegir entre diferentes

materias, minimizando la supervisión entre proyectos de grupo y permitiendo que elaboren su propia estrategia de aprendizaje (que posteriormente puede ser evaluada), adicionalmente a la responsabilidad de tareas y actividades que requieran de una estrategia condicionada, hace posible alentar a los estudiantes menos motivados, para emprender actividades de aprendizaje que generen interés propio, que con éxito, determinará que continúen educándose fuera de la escuela, mucho más allá de motivaciones externas como lo pueden ser las calificaciones

Asimismo, estudios señalan que los profesores pueden apoyar a los estudiantes a dominar las habilidades de lectura mediante tareas simples como solicitar la redacción de un poema nuevo, o solicitándoles encontrar la idea principal y las razones que dan origen a un cuento. También, con el propósito de estimular la curiosidad, los profesores pueden mostrar fragmentos de información acerca de un tema en particular, que carezca de detalles importantes o relacionar el tema con su vida personal.

En relación a la cantidad de libros que se tienen en el hogar, el análisis concluye que tener más de cien libros en casa supone una ventaja de 16.4 por ciento de probabilidad de ser un alumno sobresaliente en competencia lectora. Ante este hecho, es importante considerar que, aunque el hogar del estudiante incide de manera importante en su desempeño, no se debe de dejar de mirar los demás elementos que forman parte del contexto del estudiante, por ejemplo, las escuelas deben ofrecer variedad de títulos en la biblioteca, de esta manera, si los alumnos tienen dificultad en encontrarlos en el hogar, pueden acudir a esta por ellos.

En este mismo sentido, contar con computadoras con acceso a internet y programas educativos, debe ser un servicio generalizado en las escuelas ya que apoyan a los estudiantes a obtener información eficazmente. De la misma forma, las escuelas deben ofrecer espacios atractivos a la lectura, lugares amplios y silenciosos donde los estudiantes puedan leer o estudiar, de tal forma que los alumnos menos favorecidos con los espacios del hogar puedan acceder con mejor disposición a la lectura en estos sitios. De esta forma, las desventajas

generadas por el nivel socioeconómico de las familias pueden ser minimizadas en la escuela.

Aunque México presenta una de las menores brechas de género en habilidad lectora respecto de los países miembros de la OCDE, los resultados obtenidos señalan que, el género masculino tiene 9.1 por ciento menos probabilidad de obtener un puntaje sobresaliente en la prueba de competencia lectora, y que si los estudios de la madre son menores al bachillerato contarán con un 24.1 por ciento menos de probabilidad asociada a dicha característica.

Se observa que, la brecha de género en comprensión lectora es mayor en entornos más disruptivos y en escuelas con menor concentración de estudiantes cuyos padres tienen niveles educativos bajos. Ante esto, se sugiere que, es necesario desarrollar políticas orientadas a influir en la composición social de los centros educativos, de manera que la población de alumnos en las escuelas sea más heterogénea en los estratos socioeconómicos, el estudio señala que estas medidas previenen una alta concentración de alumnos con padres con bajo nivel educativo, o familias con niveles elevados de privación económica, y así obtener una mayor equidad en el sistema escolar que favorecería tanto a mujeres como a hombres.

Como se observó, la influencia del entorno escolar y el grupo de iguales en los estudiantes también podrían ser utilizados como una oportunidad para incrementar su rendimiento académico. Se ha observado que unas condiciones desfavorables del entorno escolar son altamente influyentes en los estudiantes y son retroalimentadas por el efecto de grupo. Sin embargo, unas condiciones favorables pueden crear un entorno positivo para el desarrollo educativo tanto de las mujeres como de los hombres.

Por tanto, se considera que se debe aplicar una estrategia “win-win” (todos ganan) que se basaría en que las mismas reformas que ayudarían a los estudiantes en el logro educativo también ayudarían a las estudiantes. No se trataría de crear reformas orientadas únicamente a reducir la brecha de género o favorecer a algún sexo, sino reformas que mejoren el entorno escolar, que aumenten el nivel de expectativas y motiven a todos los alumnos a invertir en sus estudios.

## Bibliografía

- Arnason, A. (2006). *Encouraging intrinsic motivation in the middle school reading classroom*. Acureyi University Press.1-35.
- Anagnostopoulou, K., Hatzinikita, V., & Christidou, V. (2010). Assessed students' competencies in the Greek school framework and the PISA survey. *Review of Science, Mathematics and ICT Education*, 4(2), 43-61.
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge University press.
- Brozo, W. G., Shiel, G., & Topping, K. (2007). Engagement in reading: Lessons learned from three PISA countries. *Journal of Adolescent & Adult Literacy*, 51(4), 304-315.
- Brewster, C., & Fager, J. (2000). *Increasing student engagement and motivation: From time-on-task to homework*. Portland, OR: Northwest Regional Educational Laboratory.
- Chiu, M. y McBride-Chang, C. (2006). Gender, context, and reading: a comparison of students in 43 countries. *Scientific Studies of Reading*, 10(4), 331-362.
- DiPrete, Thomas. A. y Buchmann, Claudia (2013). *The Rise of Women: The Growing Gender Gap in Education and what it Means for American Schools*. New York: Russell Sage Foundation.
- Geske, A., & Ozola, A. (2008). Factors influencing reading literacy at the primary school level. *Problems of Education in the 21st Century*, 6, 71, 77.
- Greene, W. H., & Hensher, D. A. (2009). *Modelling ordered choices*. Department of Economics, Stern School of Business, New York University, New York, NY, 10012.
- Harvey, A. C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica: Journal of the Econometric Society*, 461-465.
- Heckman, J. J., & Willis, R. J. (1977). A beta-logistic model for the analysis of sequential labor force participation by married women. *Journal of Political Economy*, 85(1), 27-58.
- Hvistendahl, R., & Roe, A. (2004). The literacy achievement of Norwegian minority students. *Scandinavian Journal of Educational Research*, 48(3), 307-324.

- Julià, A. (2016). "School Context and Gender Inequalities in Reading Achievement". *Revista Española de Investigaciones Sociológicas*, 156: 41-58.
- Kasapoglu, K (2014). A Logistic Regression Analysis of Turkey's 15-year-olds' Scoring above the OECD Average on the PISA'09 Reading Assessment. *Educational Sciences: Theory and Practice*, 14(2), 649-667.
- Kayır, Ç. G., & Erdoğan, M. (2015). The Variation in Turkish Students' Reading Skills based on PISA 2009: The Effects of Socio-Economic and Classroom-Related Factors. *International Online Journal of Educational Sciences*, 7(4).
- Ma, X. (2008). Within-school gender gaps in reading, mathematics, and science literacy. *Comparative Education Review*, 52(3), 437-460.
- McFadden, D. L. (1976). Quantal choice analysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4* (pp. 363-390). NBER.
- Marks, G. Accounting for the gender gaps in student performance in reading and mathematics: evidence from 31 countries *Oxford Review of Education* vol. 34, No. 1, february 2008, pp. 89–109.
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2006a) El programa PISA de la OCDE, ¿Qué es y para que sirve? 2006. Disponible en: <https://www.oecd.org/pisa/39730818.pdf>
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2009b) Perspectivas económicas para América Latina. Disponible en [http://www.oecd-ilibrary.org/development/perspectivas-economicas-de-america-latina-2009\\_leo-2009-es](http://www.oecd-ilibrary.org/development/perspectivas-economicas-de-america-latina-2009_leo-2009-es)
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2009c). *PISA 2009 assessment framework—Key competencies in reading, mathematics, and science*. Disponible en <http://www.oecd.org/dataoecd/11/40/44455820.pdf>.
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2009d). *Manual SPSS para PISA*. Segunda Edición. Disponible en: <http://browse.oecdbookshop.org/oecd/pdfs/free/9809031e.pdf>.
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2009e). *Cuestionario de Estudiante*. Disponible en <http://PISA2009.acer.edu.au/downloads.php>.

- Rasmussen, J. B. (2003). Reading literacy performance in Norway: Current practice and critical factors. *European Journal of Education, 38*(4), 427-443.
- Shiel, G., & Eivers, E. (2009). International comparisons of reading literacy: What can they tell us? *Cambridge Journal of Education, 39*(3), 345-360.
- Smith, J. K., Smith, L. F., Gilmore, A., & Jameson, M. (2011). Students' self-perception of reading ability, enjoyment of reading and reading achievement. *Learning and Individual*
- Topping, K. (2006). PISA/PIRLS data on reading achievement: Transfer into international policy and practice. *The Reading Teacher, 59*(6), 588-590.
- Twist, L., Gnaldi, M., Schagen, I., & Morrison, J. (2004). Good readers but at a cost? Attitudes to reading in England. *Journal of Research in Reading, 27*(4), 387-400.
- Wald, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *The Annals of Mathematical Statistics, 10*(4), 299-326.
- Williams, T., Williams, K., Kastberg, D., & Jocelyn, L. (2005). Achievement and affect in OECD nations. *Oxford Review of Education, 31*(4), 517-545.
- Zuzovsky, R. (2010). The impact of socioeconomic versus linguistic factors on achievement gaps between Hebrew-speaking and Arabic-speaking students in Israel in reading literacy and in mathematics and science achievements. *Studies in Educational Evaluation, 36*, 153-161.